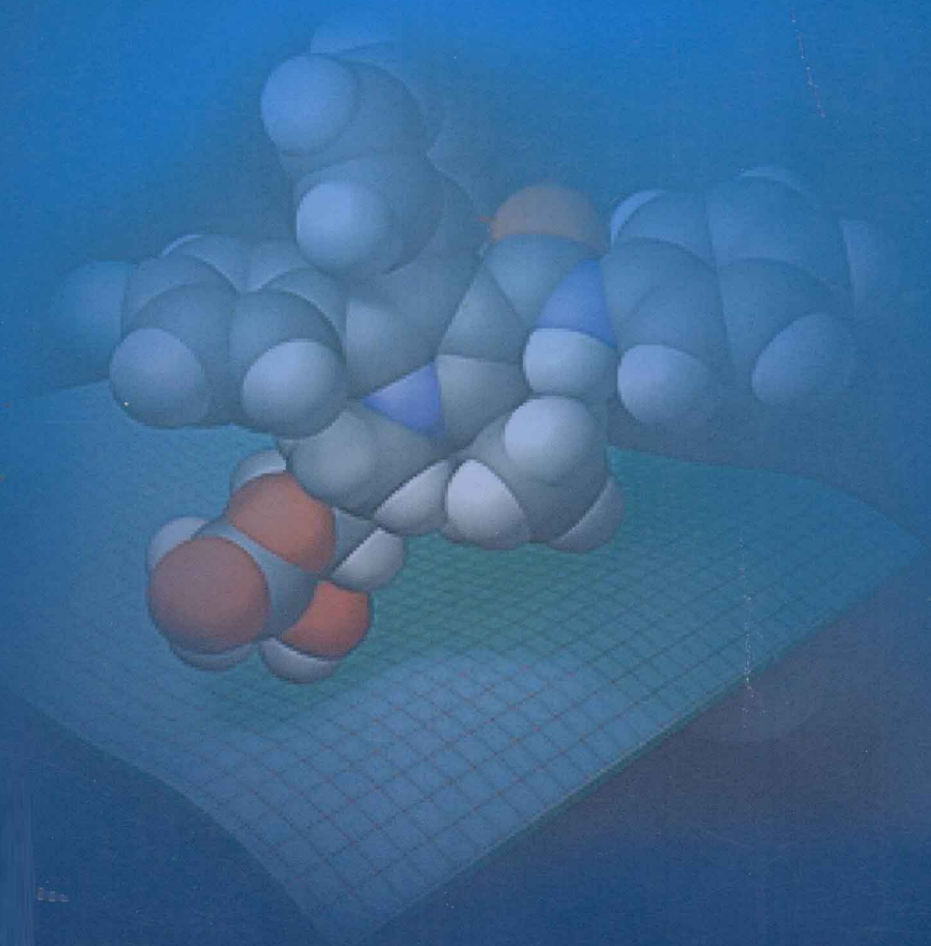


U A X U E X I N X I X U E

化学信息学

丁言斌 王 晶 编著



中国矿业大学出版社

China University of Mining and Technology Press

化 学 信 息 学

丁言斌 王 晶 编著

中国矿业大学出版社

内 容 提 要

本书是一本介绍化学信息学及其应用的技术型图书。主要介绍了文献检索的一般知识以及各类文献的检索方法。以实用性为特点,强调实践并侧重于计算机技术与计算机网络技术的应用。全书主要内容有:文献基础知识,计算机检索基础与因特网的使用,美国化学文摘,科学引文索引,中文全文数据库,英文全文数据库,电子图书及检索,工具书的利用,专利文献的查阅等。书中对各类文献的网上检索也作了介绍,并提供相关网络资源地址,以方便读者上网检索。

本书可作为高等院校化学、化工、材料、制药等专业本科高年级学生及研究生“化学信息学”课教材,亦可作为相关领域的广大科研工作者的参考书和应用工具书。

图书在版编目(CIP)数据

化学信息学/丁言斌,王晶编著. 徐州:中国矿业大学出版社,2011.10

ISBN 978 - 7 - 5646 - 1153 - 8

I. ①化… II. ①丁…②王… III. ①化学—机器检索:情报检索 IV. ①G252.7

中国版本图书馆 CIP 数据核字(2011)第 145655 号

书 名 化学信息学
编 著 丁言斌 王 晶
责任编辑 周 红 褚建萍
出版发行 中国矿业大学出版社有限责任公司
(江苏省徐州市解放南路 邮编 221008)
营销热线 (0516)83885307 83884995
出版服务 (0516)83885767 83884920
网 址 <http://www.cumtp.com> E-mail: cumtpvip@cumtp.com
印 刷 徐州中矿大印发科技有限公司
开 本 787×1092 1/16 印张 15.5 字数 397 千字
版次印次 2011 年 10 月第 1 版 2011 年 10 月第 1 次印刷
定 价 28.00 元

(图书出现印装质量问题,本社负责调换)

前 言

人类社会已经进入信息时代,科技文献宝库是人类共享的资源。随着科学技术的迅猛发展,信息社会的挑战使得信息、情报工作已经成为全方位的社会需求,而科技文献资料则是信息、情报的重要来源。对于信息利用程度的高低已经成为一个国家综合实力的重要体现。因此,培养具有较强的情报意识、善于运用前人的优秀科学技术成果、能进行开拓创新的人才是社会的要求,更是学校人才培养的方向。面向现代化、面向世界、面向未来的智能型人才的培养,更使得以培养学生的信息意识、提高独立获取知识和科学研究能力为目的的化学信息学课,成为高等教育中不可缺少的一门科学方法课。

学好“化学信息学”课的关键是要多动手、多动脑、多动口,贵在实践。如果有一检索课题,仅是通过情报站的检索人员检索,由于专业上的问题,往往得不到很满意的结果。其次,各工具书有各自的特点,所以要用多种途径进行查找,不可能有哪一本工具书能够解决检索中遇到的所有难题。这就要向有经验的从事文献检索的教师、研究人员、图书馆工作人员请教。

全书共分九章,各章节编写分工如下:王晶编写第一章、第二章、第五章;丁言斌编写第三章、第四章、第六章、第七章、第八章、第九章。

由于编写《化学信息学》的难度非常大,编者的经验与学识均感不足,书中难免有许多不当之处,真诚欢迎读者提出宝贵意见。

丁言斌

2011年5月于淮北

目 录

第一章 绪论	1
第一节 文献基本知识	1
第二节 文献检索	5
第三节 检索的策略、方法与步骤	8
第四节 化学文献总览	12
第二章 计算机检索基础与因特网的使用	20
第一节 计算机检索基础知识	20
第二节 因特网的使用	24
第三节 网络资源的综合利用	27
第三章 美国化学文摘(CA)	33
第一节 CA 一般情况介绍	33
第二节 CAS 产品和服务系列	38
第三节 CA 文摘及其组织	42
第四节 CA 检索实例	52
第四章 科学引文索引(SCI)	77
第一节 SCI 一般情况介绍	77
第二节 SCI 产品和服务系列	78
第三节 SCI 的用途	82
第四节 SCI 检索	85
第五节 SCI 分析工具	93
第五章 中文全文数据库	99
第一节 中国国家知识基础设施工程	99
第二节 万方数据资源系统	115
第三节 中文期刊数据库(重庆维普)	122
第六章 英文全文数据库	130
第一节 Science	130
第二节 Nature	140
第三节 Elsevier Science Direct	146

第四节	Wiley InterScience	157
第五节	Springer	165
第六节	American Chemical Society	175
第七节	RSC	186
第七章	电子图书	197
第一节	电子图书的特点	197
第二节	电子图书的主流格式	198
第三节	电子图书的检索	198
第八章	工具书的利用	209
第一节	工具书的特点	209
第二节	数值数据的检索	210
第三节	化学品原料和中间体的信息获得	216
第四节	有关技术和工业应用的工具书	217
第五节	有关化学物质安全、毒性等方面的信息	218
第六节	人物和机构	219
第九章	专利文献的查阅	221
第一节	专利基础知识	221
第二节	国际专利分类表	225
第三节	专利的查阅	226
第四节	其他专利文献	236
参考文献		239

第一章 绪 论

第一节 文献基本知识

文献检索又称情报检索、信息检索。文献检索指检索文献中的情报,情报检索指检索载有情报的文献,信息检索指从信息载体中检索所需信息。

一、文献的定义

文献的定义是:文献是记录有知识的一切载体。它们是人类进步和发展的记录和积累。国际标准化组织的定义:在存储、检索、利用或传递记录信息的过程中,可作为一个单元处理的,在载体内、载体上或依附载体而存储有信息或数据的载体。具体地讲,文献是指以文字、符号或图形等方式记录人类活动或知识的一种信息载体。它将人类知识用文字、符号、图形、数字、代码及音频、视频信号等记录方式在甲骨、竹帛、纸张、胶卷、磁带、光盘等载体上记录下来而形成的。知识、载体和记录是构成文献的三个基本要素。知识是文献的内容,载体是文献的外在形式,记录则是两者联系的手段。科技文献就是记录科技知识内容信息的集合体。

文献的属性有以下几个方面。

1. 知识信息性

这是文献的本质属性。任何文献记录或传递的都有一定的知识信息。离开知识信息文献就不存在。

2. 客观物质性

文献记录的知识信息只有通过各种记录手段、记录在一定的载体上才能保存和传递。

3. 人工记录性

知识信息是人们通过各种方式记录在载体上的。

4. 动态发展性

文献并不是静止的,随着人类记录水平的提高,文献的数量在增长,载体形式在多样化,同时一部分文献在老化,失去价值。

随着科学技术的飞速发展,文献信息量飞速增长。原始文献的产生是无序的,为了从数量不断增长的无序文献中获得有用信息,便产生了文献检索工具。科技人员获得文献信息的手段是多种多样的。他们不仅利用检索工具来获得自己感兴趣的信息,还通过浏览和阅读最新科技文献,进行各种各样的学术交流和私人通讯来获取感兴趣的信息。

二、文献的分类

由于科技文献数量庞大、品种繁多、特色各异,它们所记载的信息内容也各有侧重。因

此,首先应了解一下不同科技文献的特点,把文献按其形式和类型归类,能够从内在规律来进一步认识文献。这对深入了解文献、熟悉文献、进行科技文献检索会有很大的帮助。

科技文献的分类通常有两种方法。一种是文献情报工作者按文献的性质对科技文献进行划分,他们把科技文献划分成一次文献、二次文献、三次文献和零次文献。另一种则是按出版物类型对科技文献进行划分。了解这两种划分方法可对科技文献概貌有一总体上的认识。

(一) 按文献性质划分

1. 一次文献

直接记录原始研究结果的文献资料称为一次文献,习惯上也称原始文献或第一手资料。期刊上发表的论文大多是一次文献。此外如科技报告、会议论文、专利说明书、学位论文、技术标准等等,只要是记载原始研究结果的论文均属一次文献。优点:具有新颖性、创造性(是人类创造性劳动成果)和先进性(内容上是新成果、新发明);叙述具体,有观点、有材料、有数据,具有参考、借鉴和利用价值。缺点:数量大,分布散,难以查找。

2. 二次文献

对一次文献进行加工整理后的产物,目的是使无序的一次文献(也包括三次文献)可以被检索。这类文献有文摘、题录、索引等检索性出版物。此外,对一次文献所含信息进行压缩和归纳整理,并以简洁、易读和易检索形式编辑的出版物也属二次文献,如百科全书和手册等。形式上有卡片式、期刊式、书本式。二次文献是通常认为的检索工具。CA(Cheical Abstracts)和SCI(Science Citation Index)就是最典型的二次文献。特点:汇集性,即汇集某一特定范围文献,在此基础上加工、整理而成完整反映某一情报部门、某学科、某专题文献;工具性,即按一定方式汇集特定范围文献,并以特定方式揭示文献外部特征和内部特征,以供人们查找;系统性,即对一次文献进行系统编排,使之有序化。

3. 三次文献

按某专题对一次文献进行综合整理所形成的文献资料,如专题文献总结(罗列性的)、专题文献述评(对罗列的文献进行分析和评价,有的述评作者是这一领域的权威学者)等。三次文献散见于期刊和一些专门的综述性刊物中,许多图书也属三次文献。许多人把文献检索的注意力集中于二次文献,实际上,能在文献检索中获得1~2篇总结性或述评性的三次文献帮助更大。三次文献也可通过二次文献检索得到。有经验的文献检索者都知道,在二次文献的检索结果中首先阅读三次文献,而三次文献中的参考文献引证又是扩大检索的重要途径。特点:综合性,即将大量分散资料进行综合、分析、评价、筛选,可横向综合、纵向综合利用;价值性,即在大量文献基础上提炼而成,比较可靠;针对性,即针对某一特定目的进行情报研究产生成果。

4. 零次文献

一种尚未形成文献的信息,为科技人员之间的思想交流,是非文献系统内的信息传递。严格来讲,它不能称文献,称信息更适宜。在会议中的非正式讨论、私人通信中的讨论均可视做零次文献信息。

在科学研究中,十分重视零次文献信息的获得。从原始文献获得信息,通常有半年以上的滞后,通过二次文献获得信息,滞后更多。通过会议获得信息滞后少,而会议中的讨论可获得科技人员的新思想。现在,随着通讯网络的发展,Internet的各种服务功能为零次信息

的获得增添了新的有效渠道。

(二) 按出版物类型划分

1. 期刊

期刊(Periodicals)亦称杂志(Journal 或 Magazine),为定期或不定期的连续出版物,因此又称之为连续出版物(Serials)。期刊分为:原始性期刊,如 JACS(Journal of the American Chemical Society)、JCS(Journal of Chemical Society)、化学学报等,检索性期刊,如 CA(Chemical Abstracts)、中国化工文摘等,述评性期刊,如 Chemical Reviews、Accounts Chemical Research、化学进展等,消息性期刊,如 C&E News、Chemical Marketing Reports,科普性期刊,如 Scientific American 等。

期刊中包含一、二、三次文献。CA 中摘录的期刊有 18 000 多种。文献的三大支柱是期刊、专利文献和科技图书。按出版周期分为年刊、月刊、周刊、双月刊、季刊。按学术水平和编辑部的资格等标准分为市级、省级、全国、国际等级别。

在原始性期刊中发表的论文通常有三种:论文全文(Paper)、简报或摘记(Note)和通讯(Communication or letter)。

论文全文通常有文摘(Abstract)、前言(Introduction)、实验(Experiment)、结论(Conclusion)和讨论(Discuss)等部分。简报的内容不如论文有分量,重点在实验部分,前言、结果和讨论都很简短。通讯对那些即将发表的论文中特别重要的问题予以抢先报道,因而反映新思想。

2. 图书

图书(Book)是一种重要的科技文献。它大多是对已发表的科技成果、生产技术知识和经验的概括论述,而不直接记录科研、生产的成果,一般来说它是经过总结和重新组织的三次文献,是论述或介绍某一学科或领域知识的出版物。它是作者对已发表的科研成果及其知识体系进行选择、融会贯通、概括和总结的产物。要了解和系统地掌握一门学科、一个专业、一种方法,往往都利用图书获得。它的特点是:

- ① 内容成熟、全面、系统;
- ② 出版形式较固定;
- ③ 但出版周期较长,传播速度慢。

图书可分为阅读类图书,如专著、教科书、科普读物、论文集;参考类图书,也称参考工具书,如字典、词典、百科全书、年鉴、手册、组织机构指南、企业名录、图谱、地图等。

3. 专利

通常所说的专利文献,是指向专利局申请专利时,呈交的一份详细说明发明的目的、构成及效果的书面技术文件——专利说明书。专利说明书形式有公开专利和批准专利两种。如日本、德国、美国专利有所不同。广义的专利文献还包括专利公报(摘要)及专利的各种检索工具(目录、索引、分类表及指南等)。虽然专利说明书本身是作为专利局审批专利的法律文件而存在的,但从世界科技发展的趋势和它的社会应用性看,越来越显示出专利文献是一项极重要的情报资源。它具有下列特点:

- ① 数量庞大、内容丰富,具有新颖性、独创性,属一次文献。化学专利包括化合物的合成、工艺、用途、各种配方等内容。
- ② 具有较强的实用性和可靠性。

③ 它一般只论及局部而难以看清全貌,它不提供准确的技术数据,很少说明技术原理等。

4. 科技报告

科技报告是研究或承包单位向为其提供经费的部门反映研究过程、阶段进展情况的备忘录或报告,以及工作最终成果的正式技术文件。科技报告的类型包括技术报告书、技术备忘录、札记、通报等等。最著名的科技报告有 PB(Office of the Publication Board,美国商务部出版局)、AD(Armed Services Technical Information Agency,美国武装部队技术情报局)、NASA(National Aeronautics and Space Administration,美国国家航空和宇航局)、DOE(Department of Energy,美国能源部,原为 US AEC 报告,ERDA 报告)等,这些被称为四大报告。我国科研成果的统一登记和报道工作,是从 1963 年正式开始的,由国家科学技术委员会负责发表成果公报和出版研究成果报告。从 1971 年起,改由中国科学技术情报研究所出版,报告名称统一改为“科学技术研究成果报告”,分为内部、秘密、绝密三个级别,由内部控制使用。

5. 会议文献

会议文献就是学术会议上所提出的论文。发表过的论文不能参加学术会议,而会议上提交的论文还可正式发表,因此科学研究的新结果往往在会议上先发表。科技会议文献可分为会前文献和会后文献两种。会前文献通常是提交会议的论文,往往以预印本、论文摘要等在会议上传播。参加会议不仅能得到会前文献,更重要的是可以通过会议中的交流了解正在进行中的科学研究的动向。美国化学会每年两次年会,会议的论文摘要(Abstracts of National Meeting of American Chemical Society)就是很有价值的文献资料,具有相当高的引文影响因子。会前文献大多为非正式出版物,通常只有会议代表能得到,不易收集。会后文献是主要的会议文献,通常为专家审定后录用的论文的正式出版物。会后文献常以图书、期刊中的专集等各种形式出版。例如,ACS(The American Chemical Society,美国化学学会)出版的丛书(ACS Symposium Series)是著名的会后文献。

6. 学位论文

学位论文是高等学校或科研单位的研究生为取得某种学位而撰写的研究性论文。学位论文在英国称为 Thesis,在美国称为 Dissertation。学位论文从内容上看,一般都参考了大量国内外文献资料,对研究的课题有详细的记述和论证。一些优秀的学术论文的主要部分通常会在期刊上发表。

7. 技术标准

技术标准主要是对工农业产品和工程建设的质量、规格及其检验方法等方面所作的技术规定。它是从事生产、建设的一个共同的技术依据,是一种规章性的文献,有一定的法律约束力。在化学领域的标准中,对于各种化学品的规格均有详细规定,这些规定是产品的质量检验标准。

国际上最重要的标准化组织为 ISO(国际标准化组织 International Organization for Standardization)。我国的标准体系把中国的标准定为国家标准、部颁标准和行业标准。

8. 产品资料

产品资料主要是产品目录或产品说明书。在化学领域,许多公司(如 Merck、Sigma-Aldrich 等试剂公司)都提供自己的化学品目录。

9. 电子出版物

电子出版物,指以数字代码方式将图文声像等信息存储在磁、光、电介质上,通过计算机或者具有类似功能的设备阅读使用,用以表达思想、普及知识和积累文化,并可复制发行的大众传播媒体。媒体形态包括软磁盘(FD)、只读光盘(CD-ROM)、交互式光盘(CD-I)、图文光盘(CD-G)、照片光盘(Photo-CD)、集成电路卡(IC Card)和新闻出版社认定的其他媒体形态。

随着信息技术的发展,电子出版物和联机系统在文献中的地位越来越重要,发展十分迅速。最初的电子出版物产品的载体是磁带,主要是二次文献,它们是作为书本式检索工具的副产品而出版发行的。现在,二次文献的编辑出版机构已经把机读产品作为竞争和发展的主要目标。当前,电子出版物的主要形式是光盘,光盘出版物已经从单纯的检索型发展到全文型和多媒体型。最近,电子出版物已经发展到直接在网络上出版。商用联机检索系统(如Dialog、STN等)已有20多年的历史,而Internet中众多的数据库更是大大丰富了联机系统的内容。联机系统中除了检索性的文献数据库、数值数据库、事实数据库以外,还有全文数据库。全文数据库包括全文图像数据库和全文文本数据库。

第二节 文献检索

一、文献检索工具

在浩瀚的文献海洋中,要找到感兴趣的文献,就必须借助于检索工具。信息检索工具就是用来报道、存储和查找文献信息的检索工具,是在一次文献基础上加工、提炼、编辑而成的二次文献,能使用户广、快、准、省(省时、省力、省钱)地获取有关文献信息。这里说的存储并不是存储文献的全文,而是把有情报价值的文献先选出来,再把其中的某些特征著录下来,成为一条条文献线索(文献标识),最后将这些文献标识有序化,排列成有条不紊的整体。这样文献标识的有机集合就脱离了文献本身,而单独成为查找文献的一部检索工具。对计算机文献检索而言,则成为一个计算机文献检索系统。文献特征包括外表特征:题名、著者、来源、卷期、页次、年月、号码、文种等;内容特征:分类号、主题词、文摘;检索标识:描述文献外部特征和内容特征的专门用于信息检索的词、词组或代码。

(一) 检索工具必须具备的基本条件

- ① 收录有足够的文献,且这些文献都含有一定的情报量;
- ② 对所收录文献的各种外部特征和内容特征有详细描述,使之成为一条完整的文献著录款目;
- ③ 每条款目上方标有检索用标识:分类号、主题词、著者姓名、文献序号;
- ④ 全部文献都按标识的顺序,系统、科学地排列,科学组成有机整体;
- ⑤ 提供多种检索途径,提供多方面、多角度查询。

(二) 检索工具的职能

检索工具一般应有多种索引,如著者索引、分类索引、主题索引等。其职能有以下三个方面。

1. 报道职能

检索工具可以揭示某一时期某一范围的科技文献的发展状况。

2. 存贮职能

检索工具能把有关文献的内容特征和外部特征著录下来,成为可以识别的一条条文献,并按一定的次序组织起来,以便于查找信息。

3. 检索职能

提供一定的检索手段,使人们按一定的方法,及时、准确、全面地找到所需的信息。

二、文献检索工具分类

文献检索工具可分为手工检索和计算机检索。

(一) 手工检索

手工检索工具的类型按著录方式分为以下几种。

1. 目录

目录是以单独出版物为报道单位,揭示外表特征的检索工具。通常以整本的图书、期刊等作为报道单元,注明具体的出版或收藏单位。它只描述文献的外部特征,不涉及书中的具体章节或期刊中的具体文章。

2. 题录

题录是揭示具有重要检索意义的内容标识,并注明文献条目线索的检索工具。以单篇文献作为报道单位,揭示外表特征的检索工具。它按照论文的篇名顺序排列,具有快和全的特点。虽然它提示文献的程度比目录要深,但仍只是外部特征。

3. 索引

索引是将文献中某些重要的、具有检索意义的内容特征标识或外部特征标识按某种顺序排列并注明文献条目线索的检索工具。主要指图书后所附的“索引”或检索工具后所附的“辅助索引”,是提示各种文献外表特征或内容特征的系统化记载工具。索引作用如同钥匙一样,将文献内容中有价值的情报资料按分类字顺的序列显示出来,使读者迅速获取文献中知识单元。索引将图书、期刊中所载文章的篇名、著者、主题、人名、地名等著录出,标明出处,并按一定顺序排出,以供人们查找使用。

4. 文摘

文摘是以单篇文献作为报道单位,揭示外表特征和内容特征的检索工具。文摘以精练的语言把文献信息的重要内容、学术观点、数据及结构准确地摘录下来,并按一定的著录规则与排列方式编排起来,供读者查阅。文摘其实就是题录加内容摘要。如1830年德国创刊的《药学总览》(后改名为《化学总览》)、1907年美国创刊的《化学文摘》等都是世界上较早的文摘刊物。

文摘按篇报道、按类编排,一般附各种索引。著录项目有:篇名、著者、出处、内容摘要。通过阅读文摘很快能掌握文献的基本内容,从中获得专业最新的发展动向和研究成果,比题录的检索功能更强。

根据对文献内容揭示深度或报道的详细程度,文摘又分为以下几种:

① 指示性文摘,又称简介或提要。它是以最简短的语言文字说明文献主题范围、研究方法、结论、用途等,不涉及具体的技术内容。实际上是对题目的补充说明,一般在100字左

右,概括性强。它的主要作用是帮助读者判断其原文是否是自己所需要的。

② 报道性文摘,又称叙述性文摘。它是以原文为基础浓缩写成的摘要。以揭示原文的主题和创造性部分为宗旨,如论文的论点、目的、结果、方法、手段,同时还包括数据、公式、图表等。篇幅一般在 500 字左右,重要文章可达 1 000 字。这种文摘可以代替阅读全文,反映的内容具体、客观、不带任何评论,如美国的《化学文摘》。

③ 评论性文摘。除浓缩原文内容外,还包括文摘员的分析和见解,如美国的《数学评论》。

④ 统计性文摘,又称表格式摘要。它是用表格形式摘录原文数据,适用于统计学、热物理学等。

(二) 计算机检索

随着科技的发展,手工检索的局限性越来越大,计算机检索能做到全面、准确、实时地获得所需的信息。计算机情报检索就是指利用计算机进行文献信息检索的过程。计算机检索包括光盘数据库检索、网络数据库检索和互联网信息检索。计算机情报检索系统必须具备以下条件。

1. 物质条件

① 计算机。它是检索系统的核心部分,包括硬件和软件。

② 通信网络。通信网络是沟通检索终端和检索系统主机的桥梁,其作用是确保信息传递的畅通。原来联机检索系统采用 TELENT 公共数据网连接,现在已发展到采用光缆、通信卫星等多种连接手段并举的阶段,通信速度有了极大的提高。主机在网络上有端口,检索用户通过公共电话线或专用线路与本地的数据通信网络相连,然后拨号,通过卫星通信与主机联机。

③ 检索终端。它是实现用户与检索系统传递信息进行“人机对话”的装置,包括计算机、电话线、调制解调器或 ISDN、ADSL 等。

④ 数据库。它是计算机检索系统中的信息源,是一系列信息记录的集合。它存储在计算机的磁带、磁盘或光盘上,借助于数据库的管理软件技术和检索系统,对于不同的信息需求,按不同目的提供各种组合信息,以满足读者的需求。近年来数据库的发展十分迅速。全世界的数据库的数量每年递增 10% 左右,几乎覆盖所有学科专业。数据库中的文档形式包括顺排文档(又称主文档)和倒排文档。前者是完整的顺序记录,后者是将主文档中的可检属性抽出来重排而生成的数据记录。

2. 人员条件

计算机检索的效果与检索人员的素质有密切的关系。人员素质包括对课题的了解程度,对检索系统的掌握程度,对检索语言、检索策略以及外语的掌握程度。

计算机检索方式按用户界面与操作方式可分为三类:命令式检索、菜单式检索和超文本检索。

① 命令式检索。是使用一些特定的操作命令(包括指令)来实施检索。适用于有经验的检索人员。

各命令的综合应用可以精确地表达检索提问,灵活地进行各种方式的检索命令,可以简捷、快速地得到比较理想的检索效果。一些大型的检索系统都使用命令检索方式。

② 菜单式检索。是一种方便易掌握的检索方式,称“easy-to-easy”检索。普通用户只要

根据菜单的指引,通过确定适当的选项和功能便能一步一步地完成检索。对于不同数据库用户只要根据提示就可检索,不需要事先培训。光盘数据库多采用菜单检索方式,其操作步骤多,检索时间长,检索功能、精度也不如命令方式。例如全文数据库。

③ 超文本检索。向用户提供更友好的人机交互界面。超文本的内容排列是非线性的,它按知识(信息)单元及其关系建立起知识结构网络,操作时用鼠标点击相关的知识单元,检索便可追踪下去,进入下面各层菜单。这种检索方式常用在多媒体电子出版物中,这类出版物不光采用超文本,而且常采用超媒体,提供文本和图像界面,Internet 上的 www 就是典型的例子。

表 1-1 列出了几种检索工具,这几种检索工具并不是检索化学文献的必需的工具。化学文献检索涉及的检索工具很多,这将在以后各章分别介绍。

表 1-1 各类文献的专门检索工具

文献类型	印刷本检索工具	机读检索工具
专利文献	World Patent Index	World Patent Index Web of Science
科技报告	Government Report Announcement and Index	National Technical Information Service(简称 NTIS)
学位论文	Dissertation Abstracts	ProQuest Digital Dissertations
会议论文	Index to Scientific and Technical Proceedings	CNKI ISTP

第三节 检索的策略、方法与步骤

一、科研人员的文献需求

(一) 情报性需求

搞科研必须及时了解世界上某科研领域的进展以及学科的前沿和热点。科研人员期望跟踪一些最有影响的权威杂志和消息性杂志,从中获取信息。

(二) 调研性需求

当一个课题开题时,必须要进行文献调研,以确定课题的新颖性和可行性,避免重复人家已经做过的研究工作,避免重复已取得专利保护的成果的开发。通过调研性的文献检索,以确定研究工作的必要性。

(三) 工作性需求

在研究工作过程中碰到问题需要查阅文献,如查一些化合物的合成方法、理化性质数据、分子结构鉴定的谱图等等。

(四) 查证性需求

当课题结题时对课题的比较和评价的查证,对发表论文的影响的查证,申请专利时的查新、疑性查证等。

(五) 竞争性需求

一些研究工作的论文必须抢先发表,否则一文不值。我国的一些最高水平并产生巨大影响的科研项目(如人工合成胰岛素、酵母丙氨酸转移核糖核酸、水稻基因等)的结果都是抢先发表的,当时的研究人员时刻关注着国际同行的研究动向。现在,面向市场经济的一些应用研究十分关心产品信息、产品的合理定价等等。

二、检索途径

(一) 检索途径的含义

检索途径也称检索入口,文献的特征是存储文献的依据,也是检索文献的依据,因此文献便构成了检索途径。

(二) 检索途径的类型

外表途径以文献外表特征标识作为检索入口的检索途径。

1. 题名途径(包括书名、篇名、刊名)

指根据文献的名称来查找文献的途径,文献名主要指书名、刊名、会议名等,文献名索引按其名称字顺排列。这类检索工具有图书书名目录、期刊刊名目录等。

2. 著者途径

指根据著者的姓名来查找文献的途径。在检索工具中有著者索引和著者目录。著者索引按著者的姓(或名)字顺排列,后给出文摘号,著者索引又分为个人著者索引和团体著者索引。由于同一著者的文章往往具有一定的逻辑联系,以著者为线索可以系统、连续地掌握他们的研究水平和研究方向,因此具有族性检索的功能。

3. 号码途径

指根据文献代码来查找文献的途径。号码有多种类型,如科技报告有报告号、标准文献有标准号、专利文献有专利号,使用这种序号索引查询很简单,但需要了解文献的号码,在事实检索中用得较多,一般情况下使用得不多。

4. 其他途径

有依据出版类型来检索的途径,有依据文献出版日期来检索的途径,有依据国别、文种来检索的途径。

(三) 文献内容特征的检索途径

在不了解文献外表特征,并不是确切查找某一指定文献,而只是查找相关文献时,可使用文献内容特征的检索途径。

1. 主题途径

以代表文献主题内容为主题词为标引检索文献,通过主题目录和主题索引进行检索的一种方法。主题词按字顺排列,建立主题词索引(关键词索引、叙词索引),后面列出文摘号。

2. 分类途径

以分类号作为检索标识,有分类目录,也有分类索引等。

大多数检索工具的正文是按分类编排的,它们的目录就是分类索引。如大型工具书CA,书前有分类目录。分类目录或分类索引是以类名分类号作为检索标识的。

3. 其他检索途径

其他检索途径有引文途径、分子式途径、化学物质途径。如CA中分子式索引、环系索

引、化学物质名称索引。

三、检索方法

(一) 分散性检索方法

分散性检索方法是在分散的文献载体中直接检索所需文献资料的方法。分散性检索方法的特点是:① 直接接触一次性文献,便于判断内容的取舍;② 时效性、检准率高;③ 系统性差,检全率低。

(二) 系统性检索方法

系统性检索方法是利用检索工具全面系统检索文献资料的方法。

1. 追溯法

追溯法,又称回溯法,它是从已有的文献后面所附的参考文献入手,逐一追查原文,再从这些原文后面的参考文献逐一追查,不断扩大检索的线索,从而获得一批相关文献的查找方法。

该方法的优点:在没有检索工具或检索工具不齐全的情况下,可以查到一批相关文献。

该方法的缺点:作者列的参考文献一般是有限的,有的与原文的关系不是很大,因此容易造成漏检和误检。

2. 常用法

常用法是利用文摘、题录或索引等检索工具来查找文献的方法,也称工具法。由于这种方法科学、省时、省力、效率高,是现在检索最常用的一种方法,所以叫常用法。

按照所查文献的时间顺序,常用法分为以下几种:

① 顺查法。是以检索课题起始年代为起点,按时间顺序由远而近地查找文献的方法。查找前需了解课题的背景,通过有关的参考工具深入了解课题的实质性内容,再选择检索工具,从问题的发生年着手查起。该方法查全率高,但工作量大,效率不高,一般在课题立项查新或成果鉴定查新时使用。

② 倒查法。是逆时间顺序,由近而远地查找文献的方法。此法多用于新课题或老技术新发展的课题,但易于漏检而影响检全率。课题对近期的状况比较重视,从新情况开始查找。

③ 抽查法。是根据课题的特点和需要,选择该课题研究发展较快、出版文献较多的年代,抽取其中若干年,再进行查找的一种方法。该方法检索时间短,有漏检的可能,只有在对课题的学科发展前景有较多了解时才能使用。

④ 分段法。是利用检索工具查出一批文献资料,再利用文献资料所附的参考文献追溯查找相关文献的一种方法。此法在检索工具缺期、缺卷时,也能连续获得所需年限以内的文献资料。

3. 综合法

综合法因综合使用追溯法和常用法而得名,又称循环法或交替法。先利用检索工具查出一定时期内的一批有用文献,然后利用这些文献后的参考文献,再以追溯法查出前一时期内的文献,如此循环交替地使用上述两种方法,直到满足要求。该方法兼有前两种方法的优点,全面而准确,适用于查阅那些过去年代内文献量较少的专业,并可弥补因检索工具不全而造成的漏检。

以上几种方法各有优缺点,实际检索课题时选用哪一种,要根据课题研究的需要以及所能利用的检索工具和检索手段。在检索工具较多的情况下,可以使用常用法;在已获得文献针对性很强的条件下,可利用追溯法获得相关性较强的文献;在获悉研究课题出版文献较多的年代即可利用抽查法。

四、检索步骤

(一) 分析研究课题

1. 分析检索课题所属学科范围

明确检索的目的和意义,弄清课题的核心含义,确定所属学科、专业范围。

2. 确定检索时间范围

根据课题的具体情况而定,课题查新就要检索近 10 年的文献,一般情况可长可短。

3. 分析文献的类型

现代科技文献种类很多,而不同类型的文献往往为不同的科技工作所需要。如从事一般性的科研、生产技术工作,需要掌握学科动态,主要利用期刊论文、会议文献、科技报告、专利文献。

例如:定型产品的设计和检验,需要利用标准文献;从事发明创造,搞技术革新,需要检索专利文献;产品的外观设计,需要利用产品资料。

(二) 选择检索工具

检索工具直接关系到检索效率和质量的高低。现在检索工具种类也很多,选择时可考虑以下几个方面的因素:

- ① 检索工具报道文献的学科专业范围;
- ② 检索工具所报道的文献类型;
- ③ 检索工具所收录文献的语种;
- ④ 检索工具提供的检索途径。

同时要从现有的检索工具情况出发。如果对本专业有哪些检索工具或检索工具报道的专业范围不太清楚,或利用介绍检索工具的工具书,如中国科学技术情报研究所的《国外科技文献检索工具书简介》及其《续编》、辽宁人民出版社出版的《科技检索工具书综录》(上、下)等。

(三) 确定检索途径和检索策略

在利用检索工具查找文献时,主要是利用检索工具的各种索引,即通过各种检索途径来查找文献线索。一般来说每种检索工具都提供多条检索途径。如果检索课题要求的是泛指性较强的文献信息,则最好选择以分类号作为检索标识的分类途径,此法系统性强、检全率高。如果要求专指性较强的文献信息,最好选择以关键词作为检索标识的主题途径,此法检准率高。如果事先已知文献题名、著者、号码(专利号、标准号、报告号)等条件,可利用相应的各种途径。但也要视选择的检索工具具体而定,有什么途径用什么途径。

检索策略就是提取检索词并确定组配关系。在检索过程中,检索策略可能要经过多次改变,如检索文献太多,就要缩小检索范围。

(四) 选择检索方法

可参照前面分散性检索方法、系统性检索方法进行选择。